

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

ФАКУЛЬТЕТ ПРИКЛАДНОЇ МАТЕМАТИКИ

Кафедра системного програмування і спеціалізованих комп'ютерних систем

«До захисту допущено»

Завідувач кафедри

(підпис) Тарасенко В.П.
(ініціали, прізвище)

“ ____ ” червня 2019 р.

**Дипломний проект
на здобуття ступеня бакалавра**

з напрямку підготовки **6.050102 «Комп'ютерна інженерія»**

на тему: Система оптичного розпізнавання символів _____

Виконав: студент IV курсу, групи КВ-53

(прізвище, ім'я, по батькові) _____
(підпис)

Керівник ст.викладач каф.СПСКС, к.т.н. Наливайчук М.В. _____
(посада, науковий ступінь, вчене звання, прізвище та ініціали) _____
(підпис)

Консультант з нормоконтролю, доц.каф.СПСКС, к.т.н. Клятченко Я.М. _____
(назва розділу) (посада, вчене звання, науковий ступінь, прізвище, ініціали) _____
(підпис)

Рецензент ст.викладач каф. ОТ ФІОТ Виноградов Юрій Миколайович _____
(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали) _____
(підпис)

Засвідчую, що у цьому дипломному
проекті немає запозичень з праць інших
авторів без відповідних посилань.

Студент _____
(підпис)

Київ – 2019 року

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

ФАКУЛЬТЕТ ПРИКЛАДНОЇ МАТЕМАТИКИ

Кафедра системного програмування і спеціалізованих комп'ютерних систем

Рівень вищої освіти – перший (бакалаврський)

Напрямок підготовки 6.050102 «Комп'ютерна інженерія»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Тарасенко В.П.
(підпис) (ініціали, прізвище)

«__» червня 2019 р.

ЗАВДАННЯ

на дипломний проект студента

Шпака Олексія Миколвйовича

1. Тема проекту: система оптичного розпізнавання символів, керівник проекту ст.викладач каф. СПСКС, к.т.н. Наливайчук М.В., затверджені наказом по університету від «22» травня 2019 р. №1330-С

2. Термін подання студентом проекту 14 червня 2019р.

3. Вихідні дані до проекту : див. технічне завдання.

4. Зміст пояснювальної записки:

Аналіз способів оптичного розпізнавання символів.

Виявлення області символу.

Штучна нейрона мережа.

Тестування системи.

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслеників, плакатів, презентацій тощо) :

ІАЛЦ.045490.005 Д1. Алгоритм зворотного розповсюдження. Схема алгоритму.

ІАЛЦ.045490.006 Д2. Система оптичного зчитування символів. Структурна схема.

ІАЛЦ.045490.007 Д3. Алгоритм аналізу розпізнаних символів. Схема алгоритму.

ІАЛЦ.045490.008 Д4. Алгоритм навчання штучної нейронної мережі. Схема алгоритму.

6. Консультанти розділів проекту

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
нормоконтроль	<u>доц.каф.СПСКС, к.т.н. Клятченко Я.М.</u>		

7. Дата видачі завдання : 13 лютого 2019р.

Календарний план

№ з/п	Назва етапів виконання дипломного проекту	Термін виконання етапів проекту	Примітка
1.	Вивчення літератури за тематикою проекту	22.04.2019	
2.	Розроблення та узгодження технічного завдання	03.05.2019	
3.	Аналіз існуючих рішень	07.05.2019	
4.	Підготовка матеріалів першого розділу дипломного проекту	11.05.2019	
5.	Підготовка матеріалів другого розділу дипломного проекту	17.05.2019	
6.	Підготовка графічної частини дипломного проекту	22.05.2019	

7.	Оформлення документації дипломного проекту	27.05.2019	
8.	Попередній огляд матеріалів диплому на кафедрі	29.05.2019	

Студент

(підпис)

Шпак О.М.

(ініціали, прізвище)

Керівник проекту

(підпис)

Наливайчук М.В.

(ініціали, прізвище)

Анотація

Пояснювальна записка має обсяг 51 сторінки, 9 ілюстрації, , 11 бібліографічних посилань.

Метою даного дипломного проекту є створення системи для оптичного зчитування тексту.

В роботі розглянуто та проаналізовано існуючі методи для знаходження області символу, а також способи для його ідентифікації найпоширенішими методами . Показано переваги та недоліки використання кожного з методів і висвітлена проблематика поставленої задачі.

Проведене тестування готового ПЗ OCR, для порівняння результатів провідних програм з написанною програмною розробкою.

Для реалізації мети дипломного проекту було розроблено: програму написану на мові C++, в якій реалізовані необхідні алгоритми.

Ключові слова: системи оптичного зчитування символів, ПЗ, C++.

Annotation

The explanatory note has a volume of 51 pages, 9 illustrations, and 11 bibliographic references.

The aim of this diploma project is to create a system for optical reading of text.

This work is considered and analyzed for all the variety of existing methods for founding symbol's area, and ways to identify it by the most common method. Advantages and disadvantages of using each method are shown and problems of given task are highlighted.

Provided software OCR was previously tested for comparing results of central programmes with written software development.

To realize the aim of the diploma project was created : a programme written on C++, which has its realized necessary algorithms.

Key words: System for optical reading of symbols, Software, C++

Поз.	Формат	ПОЗНАЧЕННЯ	НАЙМЕНУВАННЯ	Кількість аркушів	№ прим.	Примітки	
			<u>Документація загальна</u>				
			<u>Новорозроблена</u>				
	A4	ІАЛЦ.045490.002 ТЗ	Система оптичного розпізнавання символів.	4			
			Технічне завдання				
	A4	ІАЛЦ. 045490.003 ТП	Система оптичного розпізнавання символів.	1			
			Відомість технічного проекту				
	A4	ІАЛЦ. 045490.004 ПЗ	Система оптичного розпізнавання символів.	51			
			Пояснювальна записка.				
	A4	ІАЛЦ. 045490.005 Д1	Система оптичного розпізнавання символів.	1			
			Алгоритм зворотного розповсюдження				
			Схема алгоритму				
			ІАЛЦ. 045490.001 ОА				
Змін.	Арк.	№ докум.	Підпис	Дата			
Розробив		Шпак О.М.					
Перевірив		Наливайчук					
Н. контроль		Клятченко Я.М.					
Затвердив		Тарасенко В.П.					
Система оптичного розпізнавання символів Опис альбому					Літ.	Аркуш	Аркушів
						1	2
					КПП ім. Ігоря Сікорського ФПМ КВ-53		

Поз.	Формат	ПОЗНАЧЕННЯ	НАЙМЕНУВАННЯ	Кількість аркушів	№ прим.	Примітки
	A4	ІАЛЦ.045490.006 Д2	Система оптичного розпізнавання символів.	1		
			Система оптичного розпізнавання символів.			
			Схема структурна			
	A4	ІАЛЦ.045490.007 Д3	Система оптичного розпізнавання символів.	1		
			Алгоритм аналізу розпізнаних символів.			
			Схема алгоритму			
	A4	ІАЛЦ.045490.008 Д4	Система оптичного розпізнавання символів.	1		
			Алгоритм навчання штучної нейронної мережі.			
			Схема алгоритму.			
		Диск CD-ROM	Текст ПЗ. Тексти програм.	1		
			Графічний матеріал.			

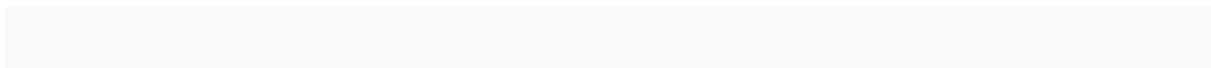
Змін.	Арк.	№ докум.	Підпис	Дата

ІАЛЦ. 045490.001 ОА

Арк. 2

ЗМІСТ

1. НАЙМЕНУВАННЯ ТА ГАЛУЗЬ РОЗРОБКИ	2
2. ПІДСТАВА ДЛЯ РОЗРОБКИ.....	2
3. ЦІЛЬ І ПРИЗНАЧЕННЯ РОБОТИ	2
4. ДЖЕРЕЛА РОЗРОБКИ	2
5. ТЕХНІЧНІ ВИМОГИ	2
5.1. Вимоги до програмного продукту, що розробляється	2
5.2. Вимоги до апаратного забезпечення	3
5.3. Вимоги до програмного та апаратного забезпечення користувача	3
6. ЕТАПИ РОЗРОБКИ	4



					ІАЛЦ. 045490.002 ТЗ						
Зм	Лист	№ докум.	Підп.	Дата	Система оптичного розпізнавання символів				Лім.	Лист	Листів
Розроб.		Шпак О.М.									
Перев.		Наливайчук								1	4
									НТУУ «КПІ ім. Ігоря Сікорського», КВ-53		
Н. контр.		Клятченко									
Затв.		Тарасенко			Технічне завдання						

1. НАЙМЕНУВАННЯ ТА ГАЛУЗЬ РОЗРОБКИ

Назва розробки: «Система оптичного розпізнавання символів».

Галузь розробки: зчитування тексту з растрового зображення

2. ПІДСТАВА ДЛЯ РОЗРОБКИ

Підставою для розробки є завдання на виконання роботи ступеня «бакалавр комп'ютерної інженерії», затверджене кафедрою системного програмування і спеціалізованих комп'ютерних систем Національного технічного університету України «Київський Політехнічний Інститут імені Ігоря Сікорського».

3. МЕТА І ПРИЗНАЧЕННЯ РОБОТИ

Метою даного проекту є розробка програмних засобів для оптичного считування тексту.

4. ДЖЕРЕЛА РОЗРОБКИ

Джерелом інформації є технічна та науково-технічна література, технічна документація, публікації у періодичних виданнях та електронні статті у мережі Інтернет.

5. ТЕХНІЧНІ ВИМОГИ

5.1. Вимоги до програмного продукту, що розробляється

- Сумісність операційною системою Windows або Linux;
- можливість обробляти нові зображення;
- можливість отримувати результат в окремому текстовому файлі;

					ІАЛЦ. 045490.002 ТЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		2

5.2. Вимоги до апаратного забезпечення

- Процесор: 2-ядерний Intel Pentium G4400;
- Оперативна пам'ять: 4 Гб;

5.3. Вимоги до програмного та апаратного забезпечення користувача

- Операційна система Windows 10;

					ІАЛЦ. 045490.002 ТЗ	Лист
						3
Зм	Лист	№ докум.	Підп.	Дата		

6. ЕТАПИ РОЗРОБКИ

№ з/п	Назва етапів виконання дипломного проекту	Термін виконання етапів
1.	Вивчення літератури за тематикою проекту	22.04.2019
2.	Розроблення та узгодження технічного завдання	03.05.2019
3.	Аналіз існуючих рішень	07.05.2019
4.	Підготовка матеріалів першого розділу дипломного проекту	11.05.2019
5.	Підготовка матеріалів другого розділу дипломного проекту	17.05.2019
6.	Підготовка графічної частини дипломного проекту	22.05.2019
7.	Оформлення документації дипломного проекту	27.05.2019
8.	Попередній огляд матеріалів диплому на кафедрі	29.05.2019

№ п/п	Формат	Позначення	Найменування	Кількість листів	Примітка
			<u>Документація загальна</u>		
			<u>Новорозроблена</u>		
1	A4	ІАЛЦ.045490.004 ПЗ	Система оптичного розпізнавання символів. Пояснювальна записка проекту.	51	
2	A4	ІАЛЦ.045490.005 Д1	Система оптичного розпізнавання символів. Алгоритм зворотного розповсюдження Схема алгоритму	1	
3	A4	ІАЛЦ.045490.006 Д2	Система оптичного розпізнавання символів. Система оптичного розпізнавання символів. Схема структурна	1	
4	A4	ІАЛЦ.045490.007 Д3	Система оптичного розпізнавання символів. Алгоритм аналізу розпізнаних символів. Схема алгоритму	1	
5	A4	ІАЛЦ.045490.008 Д4	Система оптичного розпізнавання символів. Алгоритм навчання штучної нейронної мережі. Схема алгоритму	1	
		Диск CD-ROM	Текст ПЗ. Тексти програм.	1	
			Графічний матеріал.		
			ІАЛЦ.045490.003 ВП		
Зм.	Арк	№ докум	Підпис	Дата	
Розроб.		Шпак О.М.			<div>Система оптичного розпізнавання символів.</div> <div>Відомість технічного проекту</div> <div> <div>Лім.</div> <div>Арк.</div> <div>Аркушів</div> <div>«КПІ ім. Ігоря Сікорського», ФПМ, КВ-53</div> </div>
Перевір.		Наливайчук М.В.			
Н. контр.		Клятченко Я.М.			
Затв.		Тарасенко В.П.			

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, ТЕРМІНІВ	3
ВСТУП	4
1. АНАЛІЗ СПОСОБІВ ОПТИЧНОГО РОЗПІЗНАВАННЯ СИМВОЛІВ	5
1.1. Методи розпізнавання символів	5
1.2. Ознаки символів, використовувані для автоматичного розпізнавання	6
1.3. Обґрунтування теми дипломного проекту	8
1.4. Шрифт залежні та шрифт незалежні алгоритми	9
2. ВИЯВЛЕННЯ ОБЛАСТІ СИМВОЛУ	18
2.1. Задача розпізнавання та її основні методи	18
2.2. Кроки обробки зображення	21
2.3. Сегментація	24
2.4. Шум в зображеннях	26
2.5. Математична модель зображення	30
3. ШТУЧНА НЕЙРОНА МЕРЕЖА	32
3.1. Штучна нейрона мережа та її види	32
3.2. Алгоритм роботи нейромережі	36
3.3. Навчання нейромережі	38
3.3.1. Навчання з учителем	38
3.3.2. Навчання без учителя	39
3.3.3. Навчання засноване на корекції помилок	40
3.3.4. Конку rentне навчання	41
3.3.5. Навчання Хебба	41
3.3.6. Алгоритм зворотнього поширення помилки	41
4. ТЕСТУВАННЯ СИСТЕМИ	44
ВИСНОВОК	50

					ІАЛЦ. 045490.004 ПЗ			
Зм.	Лист	№ докум.	Підп.	Дата	Система оптичного розпізнавання символів Пояснювальна записка	Літ.	Лист	Листів
Розроб.	Шпак О.М.							
Перев.	Наливайчук						1	51
Н.контр.	Клятченко					НТУУ "КПІ" ФПМ КВ-53		
Затв.	Тарасенко							

ДОДАТКИ

Додаток 1. Копії графічних матеріалів:

Додаток 2. Лістинг програмного коду

Додаток 3. Копія слайдів презентації

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		2

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАЧЕНЬ, ТЕРМІНІВ

OCR - optical character recognition

ANN - artificial neural networks

ADC - analog-to-digital conversion

NN - neural networks

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		3

ВСТУП

Для ефективного та швидкого документообігу в сучасному світі використовуються системи оптичного розпізнавання тексту (OCR). Завдання такої системи полягає в тому, щоб зробити введення документів на комп'ютер без довгої і непродуктивної "перерви". Для організації процесу розпізнавання символів використовуються безліч алгоритмів та з залученням нейронних мереж.

В останні десятиліття, завдяки використанню сучасних комп'ютерних технологій, були розроблені нові методи обробки та розпізнавання зображень, що дає можливість створювати такі системи розпізнавання, які б відповідали основним вимогам систем автоматизації документообігу. Однак створення кожного нового продукту в цій області залишається творчим завданням і вимагає додаткових досліджень у зв'язку з конкретними вимогами до дозволу, швидкості, надійності розпізнавання і ємності пам'яті.

Метою даного дипломного проекту є розробка програми розпізнавання тексту з зображення.

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		4

1. АНАЛІЗ СПОСОБІВ ОПТИЧНОГО РОЗПІЗНАВАННЯ СИМВОЛІВ

1.1. Методи розпізнавання тексту

В сучасному світі існує дуже багато різних програм розпізнавання символів, але слід зауважити, що вміння читати людиною друкований текст не високої якості все ще перевершує можливості комп'ютера. Кожен надрукований текст має деяку властивість - шрифт, яким він набирається. Відповідно, існують два типа алгоритмів розпізнавання символів: шрифт-залежних і незалежних від шрифту. Шрифт-залежні алгоритми використовують важливу інформацію про шрифт з літерами. Це означає, що програма повинна мати повний зразок тексту, надрукованого цим шрифтом. Програма вимірює і аналізує різні характеристики шрифтів і вставляє їх у свою базу еталонної характеристики.

Є три основні методи розпізнавання тексту:

- Шаблонний метод;

Методи шаблону перетворюють зображення одного символу в растр, порівнюють його з усіма шаблонами, доступними в базі даних, і вибирають шаблон з найменшим числом різних точок з вхідним зображенням. Шаблонні методи досить стійкі до дефектів зображення і мають високу швидкість обробки вхідних даних, але надійно розпізнаються тільки ті шрифти, шаблони яких відомі їм. І якщо розпізнаний шрифт дещо відрізняється від посилання, шаблонні методи можуть робити помилки навіть при обробці дуже якісних зображень.

- Структурний метод;

Структурне методологічне розпізнавання зберігає інформацію не про поточне написання символу, а про його топологію. Стандарт містить інформацію про відносне розташування окремих компонентів. Перевагою методу є стійкість до зсуву і повороту символу до невеликого кута, до різних стилістичних варіацій шрифтів. Однак, коли кут повороту десятки градусів, цей метод не можна використовувати для розпізнавання символів. При застосуванні цього методу невідомі такі ознаки, як розмір букв, які розпадалися, і навіть шрифти, на яких було надруковано. Проте головною проблемою цього методу є ідентифікація, яка містить певні дефекти.[6]

1. 2. Ознаки символів, використовувані для автоматичного розпізнавання

- Ознаковий метод.

Ознакові методи будуються на тому, що картинці ставиться у відповідність вектор ознак. Суть розпізнавання в порівнянні вектора ознак з набором прикладів векторів тієї ж розмірності. Переваги методу – простота реалізації, добра узагальнююча спроможність, висока швидкість розпізнавання. Недолік методу – велика чутливість до різних дефектів зображення. Крім того, ознакові методи мають ще один недолік — на етапі становлення ознак відбувається втрата деякої інформації про символ.

Виділення ознак проходить незалежно, тому інформація про взаємне розташування елементів символів втрачається. [1]

- Еталонні методи

Еталонні методи передбачають порівняння заданого, не розпізнаного символу з набором деяких еталонів. Для цього використовують нейронні мережі, які необхідно заздалегідь заповнити еталонами. Існують кілька алгоритмів порівняння тексту з еталоном.

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		6

Найпростіший варіант – попиксельне порівняння, однак для нього необхідні рівні по розмірам зображення, що порівнюються. Інші варіанти – накладення та накладення зі зміщенням, у яких зображення ставляться у відповідність одне одному. Методи еталонного порівняння рукописного та друкованого тексту, на перший погляд, схожі, однак суттєво відрізняється. Еталони рукописного тексту, на відміну від літер шрифту, лише грубі зразки. Це значно збільшує вірогідність помилки. [1]

У типовій системі оптичного розпізнавання символів (OCR) введені символи зчитуються та оцифровуються за допомогою оптичного сканера. Після цього кожен символ піддається локалізації і розподілу, і отримана матриця підлягає попередній обробці, тобто згладжування, фільтрація і нормалізація. В результаті попередньої обробки розрізняють характерні ознаки, після яких здійснюється класифікація.

Основні підходи та рішення:

- Алгоритм скелетизації

Це спосіб розпізнавання одиночних бінарних зображень, заснований на побудові скелетів цих зображень і виборі скелета ребер і вузлів. Потім, відповідно до співвідношення ребер, їх чисел і кількості вузлів, будується таблиця відповідності зображенням. Так, наприклад, скелет кола буде одним вузлом, скелет літери «П» - три ребра і два вузли, а ребра - 2: 2: 1. У програмуванні цей метод має кілька можливих реалізацій, більше інформацію про метод скелетонізації можна знайти нижче у розділі посилання.

- Нейрона мережа

Напрямок був дуже популярний у 60-х - 70-х роках, затоплюючи їх небагато, зменшуючись, тому що солідна кількість вимагає солідарних обтяжливих потужностей, які зазвичай відсутні на простих мобільних

платформах. Однак слід мати на увазі, що нейронні мережі іноді дають дуже цікаві результати, завдяки яким вони не мають будь-якої структури, тим більше що деякі нейронні мережі здатні розпаковувати зображення інваріантності. Наприклад, мережі, засновані на неоконетронах, можуть ідентифікувати деякі характеристики, а також їх розривати, як якщо б вони поверталися.[5]

- Інваріантні числа

З геометрії зображень можна вибрати декілька чисел, інваріантних по відношенню до розміру і повороту зображень, після чого можна зробити таблицю, що відповідає цим числам конкретного зображення (майже як у алгоритмі скелетонізації). Прикладами інваріантних чисел є число Ейлера, ексцентриситет, орієнтація (в сенсі розташування головної осі інерції відносно чого-небудь інваріантного).

- Поточное процентне порівняння з еталоном

Повинна бути якась попередня обробка, щоб отримати інваріантність по відношенню до розміру і положення, потім порівняння з підготовленою базою даних стандартів зображень - якщо збіг більше певної позначки, то ми розглядаємо зображення, яке слід розпізнати.

1. 3. Обґрунтування теми дипломного проекту

У ході повсякденної діяльності, державних структур, бізнесу, академічних установ та навчальних закладів використовують велику кількість паперових документів, більшість з них написані від руки. Великий обсяг даних і знань міститься в друковані або рукописні документи, які архівуються. Потреба в оцифрування паперових документів зростає з метою подальшої обробки їхнього змісту комп'ютерними системами.

Розпізнавання тексту можна розділити на кілька напрямів, які істотно розрізняються за своїми методами вирішення. Текст також може бути друкований та рукописний. Будь-який з них може бути додатково структурованим. Наприклад, формули можуть містити різні рівні записів, такі як верхній індекс, підлінії, спеціальні позначки для математичних дій і т.д..

Найважливішими питанням з обробки та зберігання всіх великорозмірних відсканованих даних є неспроможність пошуку в документі певної фрази. Відсканований текст не можна буде скопіювати чи змінити, тому що це набір зображень.

1.4. Шрифт залежні та шрифт незалежні алгоритми

Оптичне розпізнавання символів відбувається певного шрифту.

Недоліки такого підходу:

- алгоритм повинен заздалегідь знати шрифт, який він представляє для розпізнавання, тобто, він повинен зберігати різні характеристики в базі даних;
- Якість розпізнавання тексту, введеного довільним шрифтом, буде прямо пропорційна співвідношенню характеристик цього шрифту з шрифтами, доступні в базі даних програми. [5]

Ці фактори обмежують універсальність таких алгоритмів. Для роботи з програмою розпізнавання потрібен блок налаштування певного шрифту. Очевидно, що ця одиниця внесе свою частку помилки в інтегральній оцінці якості розпізнавання або функції. Вам доведеться встановити шрифт для користувача. Програма, заснована на алгоритмі розпізнавання символів шрифту вимагає від користувача мати спеціальні знання про шрифти взагалі, про їхні групи і відмінності один від одного, про шрифти, які друкуються. Загального способу немає дізнатися, на яких шрифтах був надрукований цей документ. З іншого боку, підхід шрифту має ту перевагу,

за якою він є активно використовувати і, мабуть, використовуватиметься в майбутньому. А саме, маючи деталізовану апріорну інформацію можливо будувати досить точні та надійні алгоритми розпізнавання. Загалом, з побудови алгоритму розпізнавання шрифтів, на відміну від без шрифту надійність розпізнавання символів інтуїтивно зрозуміла і математично виражена величина. Це значення визначається як відстань у будь-який метричний простір від представленого позначення програми в процесі навчання, до персонажа, який програма намагається визнати. Другий клас алгоритмів не залежить від кадру або не залежить від шрифту, тобто алгоритми, які не мають апріорного знання символів, що надходять до них до входу. Ці алгоритми вимірюють і аналізують різні характеристики, які властиві буквам як такі, незалежно від шрифту та абсолютного розміру, як вони друкуються. У граничному випадку для не-шрифт-залежних. Можливо, немає алгоритму навчання. В цьому випадку характеристики символів вимірюються, кодуються і розміщуються в самій програмі людиною. Однак на практиці зустрічаються рідкісні випадки, коли такий шлях можливий.

Більш загальний спосіб створення бази даних характеристик полягає у вивченні програми на зразку реальних символів. Недоліком такого підходу є нижча якість розпізнавання, ніж у алгоритму шрифту. Це пов'язано з тим, що рівень узагальнення з характерними вимірюваннями значно більші, ніж у випадку з шрифт-залежними алгоритмами.

Переваги такого підходу тісно пов'язані з його недоліками. Основні з переваг:

- універсальність. Це означає, з одного боку, можливість застосування цей підхід у випадках великої різноманітності символів - з іншого боку, через харектеристики, встановлену в них.

Підсумовуючи, такі алгоритми можуть екстраполювати накопичені знання межі навчальної вибірки, тобто послідовно розпізнавати образи,

здавалося б, далеких від тих, які були присутні у навчальній вибірці.

- продуктивність. Процес вивчення незалежних алгоритмів. Звичайно, є більш простим і інтегрованим у сенсі навчання зразок не фрагментується в різні класи. Для цього немає необхідності підтримувати в базі даних характеристики різних умов для спільного їх існування. Проявом технологічної здатності є також те, що вона часто з'являється практично повністю автоматизовані процедури навчання.

- зручність у використанні програми. У випадку, якщо програма побудована на алгоритмах без ковзання, користувачу не обов'язково знати щось про сторінку, яку він хоче ввести в програму. Це також спрощує інтерфейс користувальницької програми для відшкодування вартості опцій та діалогів навчання персоналу та управління базою даних характеристик. Таким чином, це призведе до збільшення кола користувачів, зокрема для людей з мінімальною комп'ютерною грамотністю таке програмне забезпечення буде дуже корисним.

При розпізнаванні символів досить широко використовуються штучні нейронні мережі. Алгоритми, зроблені для нейронних мереж щоб розпізнавати символи, не рідко створюються в такий спосіб: картинку символу, що є вхідним зображенням для зчитування, приводиться до певного стандартного розміру. Як правило, зображення приводиться до розміру 16x16 пікселів.

Значення яскравості в вузлах нормалізованого зображення використовуються як вхідні параметри нейронної мережі. Кількість вихідних параметрів нейронної мережі дорівнює кількості розпізнаваних символів. Результатом розпізнавання є символ, який відповідає найвищому значенню вектора джерела нейронної мережі. Підвищення надійності таких алгоритмів зазвичай пов'язано з пошуком більш інформативних вхідних ознак або з ускладненням структури нейронної мережі.

Надійність розпізнавання і необхідність програми в обчислювальних ресурсах багато в чому залежать від вибору структури і параметрів нейронної мережі. Зображення фігур зменшуються до одного розміру (28x28 пікселів). Отримане зображення подається на вхід нейронної мережі, яка має три внутрішніх рівня і 10 вузлів верхнього рівня. Нижчі шари мережі не повністю пов'язані між собою. Нижні вузли мають спільний набір шкал. Все це, згідно з розробкою розробників, повинно підвищити здатність нижчих рівнів мережі розрізняти основні функції зображень. Для збільшення здатності мережі до узагальнення та зменшення кількості необхідних обчислень і пам'яті здійснюється видалення невикористаних ваг. В результаті кількість незалежних параметрів зменшується в чотири рази. Навчання нейронної мережі здійснюється на наборі 7300 символів, тест на безліч 2000 символів. Помилки розпізнавання становлять приблизно 1% на навчальному наборі і 5% на тесті.

В якості вхідних параметрів нейронної мережі замість значень яскравості в вузлах нормалізованого растру можуть бути використані значення, що характеризують різницю в яскравості. Такі вхідні параметри дозволяють краще розподіляти межі літер. Об'єкти розпізнавання зменшуються до 16x16 пікселів. Після цього вони піддаються подальшій обробці з метою ізоляції ділянок з найбільшими варіаціями яскравості.

Одним з широко використовуваних методів підвищення точності розпізнавання є одночасне використання декількох різних модулів розпізнавання і подальша інтеграція результатів (наприклад, шляхом голосування). Дуже важливо, щоб алгоритми, що використовуються цими модулями, були якомога більш незалежними. Цього можна досягти за рахунок використання модулів розпізнавання, які використовують принципово різні алгоритми розпізнавання, а також спеціальний вибір навчальних даних.

Один з цих методів був запропонований кілька років тому і базувався на використанні трьох модулів розпізнавання (машин). Перший автомобіль навчається звичайним способом. Друга машина дізнається про символи, які були відфільтровані першою машиною, так щоб друга машина бачила суміш символів, 50% з яких були визнані першою машиною правильно і 50% неправильно. Нарешті, третя машина дізнається про символи, на яких відрізняються результати розпізнавання 1-го і 2-го автомобілів. Під час тестування до всіх трьох машин застосовуються розпізнавані символи. Оцінки, отримані на виході всіх трьох автомобілів, складаються. Символ, який отримав найбільшу загальну кількість балів, видається в результаті розпізнавання.

Як правило, алгоритм розпізнавання заснований на поділі растру з зображенням букв основних ознак і подальшого використання штучної нейронної мережі для оцінки близькості вхідного зображення з символами з заданого набору букв. Результатом роботи є набір оцінок, що відображають ступінь близькості розпізнаваного символу з символами з заданого набору символів. [9]

Набір розпізнаваних символів може включати букви і цифри. Входи для розпізнавання символів перетворюються на один розмір.

Відмінною особливістю реалізованого алгоритму є використання нейронної мережі з досить великою кількістю вхідних функцій. У вихідному зображенні виділяються основні функції, що характеризують відмінності яскравості в вузлах растру. Нейронна мережа має один внутрішній рівень, що містить 100 вузлів і є універсально пов'язаним, тобто кожен вузол внутрішнього рівня з'єднаний з усіма вхідними вузлами, і кожен вузол верхнього рівня з'єднаний з усіма вузлами внутрішнього рівня. Для зменшення кількості обчислень при розпізнаванні для кожного розпізнаваного символного зображення використовуються не всі вхідні

ознаки, а лише частина, тобто вектор вхідних параметрів нейронної мережі сильно розбавлений. [8]

Навчання нейронної мережі є звичайним способом, тобто алгоритмом, який використовується для зворотного поширення помилки. Програма навчання отримує вхідний файл із зображеннями символів. При навчанні символи з цієї бази даних є циклічними. Для кожного зображення з бази даних виділяються первинні атрибути, після чого виконуються прямі і зворотні проходи через мережу. Зміна ваги мережі під час тренування проводиться після кожного символу. Крок зміни ваги мережі є постійним.

Погано впізнавані символи розглядаються частіше, ніж інші, для прискорення та вдосконалення навчання. Використовується кеш, який зберігає важкодоступні зображення. Растри для навчання вибираються як з вхідного файлу, так і з кешу. Вибір символу з кешу відбувається з урахуванням якості його розпізнавання, тобто погано впізнавані символи вибираються частіше.

Крім того, при навчанні мережі використовується регуляризація ваги мережі, тобто вводиться їх експоненціальне вицвітання.

Якість розпізнавання залежить не тільки від алгоритмів, що використовуються програмами розпізнавання та навчання нейронної мережі, але й від того, як вивчається нейронна мережа. На якість підготовки нейронної мережі впливають наступні фактори: параметри бази з тренувальним растром, розмір, метод вибору растрів, порядок розташування растрів в основі, наявність брудних символів і помилок у розмітці.

Різні фактори оптимізації можна використовувати на різних етапах навчання:

1. Крок зміни коефіцієнтів мережі.
2. Використання регуляризації мережі.
3. Історія навчання мережі.

4. Використання додаткового шуму й перекручувань символів.
5. Момент зупинки навчання. Бажано уникати як недостатнього навчання мережі, так і перенавчання.
6. Розмір кешу поганих растрів і відносна частота вибору растрів з навчальної бази даних і з кешу поганих символів.

Параметри навчання взаємозалежні і повинні вибиратися послідовно. Наприклад, при невеликій кількості навчальних ресурсів, використання спотворень символів може призвести до поліпшення якості навчання, а зі збільшенням розміру бази даних призводить до її погіршення. Використання кеш-пам'яті поганих символів на початку тренування не має особливого сенсу. Навпаки, після декількох проходів на основі тренувальних символів, більшість символів з бази визнаються дуже високою надійністю. Зміна ваги мережі відбувається в основному за рахунок растра, що міститься в кеші поганих символів.

Регуляризація (тобто введення експоненціальної вицвітання ваги в навчання) призводить до деякого погіршення якості розпізнавання. Однак використання дуже малого коефіцієнта затухання може підвищити стабільність мережі без помітних втрат для розпізнавання.

Щоб визначити найкращий час для зупинки мережі, можна періодично перевіряти якість розпізнавання на невеликій, незалежній базі даних. [9]

Порівняння якості алгоритмів розпізнавання різних символів ускладнюється тим, що відносна величина кількості правильно розпізнаних символів значною мірою залежить від конкретної бази даних, на якій проводиться тестування. На якість розпізнавання впливають також: обсяг набору розпізнаваних персонажів, технологія викладання нейронної мережі, методологія та алгоритми виділення первинних характеристик, технологія підготовки бази даних та інші фактори.

Алгоритм може бути поліпшений шляхом пошуку більш адекватного представлення структурних ознак розпізнаваних символів. Використання

великої навчальної бази даних і збільшення пам'яті нейронної мережі також може забезпечити деяке поліпшення якості розпізнавання. Проектована система повинна працювати в режимі, близькому до реального часу, і тому розроблений алгоритм повинен бути досить швидким і в той же час мати достатню точність розпізнавання. ABBYY розробили спеціальний алгоритм MDA (multilevel document analysis, багаторівневий аналіз документа). Структура сторінки аналізується методом зверху-донизу (від складових елементів до окремих символів), а відтворення електронного документа після закінчення розпізнавання відбувається знизу-догори, проте на всіх рівнях додатково діє механізм зворотнього зв'язку.

Більшість сучасних OCR діють на трьох рівнях: символів, слів, сторінок. Однак ABBYY, відповідно до принципів IPA, ввела в FineReader ще один рівень – всього багатосторінкового документа. Перш за все це знадобилося для коректного відтворення логічної структури, яка в сучасних документах стає все складніше.[10]

Саме для цього і була розроблена ADRT (Adaptive Document Recognition Technology) – технологія аналізу і синтезу документа на логічному рівні. В кінцевому підсумку вона допомагає зробити результат роботи FineReader максимально схожим на оригінал.[10]

Серед подібних до Abby FineReader систем можна назвати:

- SimpleOCR;
- OmniPage;
- Readiris;
- CuneiForm.

Окремо слід виділити Tesseract – це система розпізнавання текстів, що у 1985-1994 рр. розроблялася Hewlett-Packard, а з 2006 є вільною і

поширюється компанією Google та дуже стрімко розвивається у останні роки.[11]

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		17

2.ВИЯВЛЕННЯ ОБЛАСТІ СИМВОЛУ

2.1. Задача розпізнавання та її основні методи

Незважаючи на те, що в даний час більшість документів складається на комп'ютерах, завдання створення повністю електронного документообігу ще далека до повної реалізації. Як правило, існуючі системи охоплюють діяльність окремих організацій, а обмін даними між організаціями здійснюється за допомогою традиційних паперових документів.

Завдання перекладу інформації з паперових на електронні носії актуальна не тільки в рамках потреб, що виникають в системах документообігу. Сучасні інформаційні технології дозволяють нам істотно спростити доступ до інформаційних ресурсів, накопичених людством, за умови, що вони будуть переведені в електронний вигляд.

Найбільш простим і швидким є сканування документів за допомогою сканерів. Результат роботи є цифрове зображення документа - графічний файл. Більш кращим, в порівнянні з графічним, є текстове представлення інформації. Цей варіант дозволяє істотно скоротити витрати на зберігання і передачу інформації, а також дозволяє реалізувати всі можливі сценарії використання і аналізу електронних документів. Тому найбільший інтерес з практичної точки зору представляє саме переклад паперових носіїв в текстовий електронний документ.

На вхід системи розпізнавання надходить растрове зображення сторінки документа. Для роботи алгоритмів розпізнавання бажано, щоб на вхід надходило зображення якомога високої якості. Якщо зображення зашумлено, різке, мають низьку контрастність, то це ускладнить завдання алгоритмів розпізнавання.

Крок сегментації дуже важливий в даному підході, тому що методики класифікації, засновані на контурному аналізі, зазнають невдачі без надійної сегментації.

Обробка слідує традиційному покроковому кроку трубопроводу, але деякі з етапів були незвичайними у своїх день, і, можливо, так і залишається. Першим кроком є зв'язок компонентного аналізу, в якому обриси компоненти зберігаються.

Загалом можна виділити три методи розпізнавання образів:

1. Метод виділення. У цьому випадку проводиться порівняння з базою даних, де представлені різні модифікації зображення для кожного об'єкта. Наприклад, для оптичного розпізнавання символів можна застосувати метод для подолання типу об'єкта під різними кутами, масштабами, переміщеннями, деформаціями тощо.

Розпізнавання звукових образів, відповідно, відбувається порівняння з деякими відомими зразками (наприклад, слово, сказане кількома людьми).

2. Другий підхід - більш глибокий аналіз характеристик зображення. У разі оптичного розпізнавання це може бути визначення різних геометричних характеристик. Зразок звуку в цьому випадку піддається частоті, амплітудному аналізу тощо.

3. Наступним методом є використання штучних нейронних мереж (АНС). Цей метод вимагає або великої кількості прикладів вивчення задач розпізнавання, або спеціальної структури нейронної мережі, що враховує специфіку цього завдання. Однак вона відрізняється більшою ефективністю та продуктивністю.

Давайте коротко розглянемо математичний формалізм розпізнавання образів. Об'єкт у розпізнаванні зображень описується набором основних характеристик (атрибутів, властивостей). Основні характеристики можуть мати різну природу: вони можуть бути взяті з упорядкованого набору типу прямій лінії або з дискретного набору (який, однак, може бути наділений структурою). Таке розуміння об'єкта узгоджується з необхідністю практичного застосування розпізнавання

образів, а також з нашим розумінням механізму сприйняття об'єктом людиною.

Дійсно, ми вважаємо, що при спостереженні об'єкта людиною інформація про неї надходить від кінцевого числа датчиків (аналізованих каналів) в мозку, і кожен датчик може відповідати відповідним характеристикам об'єкта. Окрім ознак, що відповідають нашим вимірам об'єкта, існує також відмінна ознака, або група ознак, яку ми називаємо класифікуючими ознаками, і при поясненні їх значень для даного вектора X , виконується завдання природними і штучними системами розпізнавання. Зрозуміло, що для того, щоб встановити значення цих ознак, необхідно мати інформацію про те, як відомі ознаки пов'язані з класифікацією. Інформація про це з'єднання дається у вигляді прецедентів, тобто набір описів об'єктів з відомими значеннями класифікує ознаки. І згідно з цією прецедентною практикою, необхідно побудувати вирішальне правило, яке встановить безпідставний опис об'єкта значення його класифікації атрибутів.

Таке розуміння проблеми розпізнавання образів було встановлено в науці з 1950-х років. І тоді було помічено, що таке виробництво зовсім не нове. З такою постановкою методи статистичного аналізу даних, які активно використовуються для багатьох практичних завдань, таких як, наприклад, технічна діагностика, не були дійсно добре задокументовані. Тому перші кроки визнання образів проходили під знаком статистичного підходу, який диктував основні питання.

Статистичний підхід ґрунтується на ідеї, що вихідний простір об'єктів є імовірнісним простором, а ознаки (характеристики) об'єктів представляють випадкові величини, наведені на ньому. Тоді завданням даних дослідника було висунути, чомусь, статистичну гіпотезу про розподіл ознак, а точніше про залежність класифікації ознак від інших.

Статистична гіпотеза, як правило, представляла собою параметрично заданий набір функцій розподілу ознак. Типовою і класичною статистичною гіпотезою є гіпотеза про нормальність цього розподілу (різноманіття таких гіпотез статистики придумали багато). Після формулювання гіпотези необхідно було перевірити цю гіпотезу на прецедентні дані. Цей тест полягав у виборі деякого розподілу з первинно заданого набору розподілів (параметр гіпотези розподілу) і для оцінки надійності (довірчого інтервалу) цього вибору. Насправді ця функція розподілу була відповіддю на завдання, тільки об'єкт не був класифікований однозначно, але з деякими можливостями належності до класів.[3]

Статистики розробили таке ж асимптотичне обґрунтування для таких методів. Таке обґрунтування було зроблено за такою схемою: була встановлена певна функція якості вибору розподілу (довірчого інтервалу) і було показано, що зі збільшенням кількості прецедентів наш вибір з ймовірністю переходить до 1 цей функціонал (довірчий інтервал домагався 0). Дивлячись у майбутнє, скажемо, що статистичний погляд на проблему розпізнавання виявився досить плідним не тільки з точки зору розроблених алгоритмів (включаючи кластер, дискримінантний аналіз, непараметричну регресію та ін.), але згодом приніс Вапника до створення глибокої теорії статистичного розпізнавання.

2.2. Кроки обробки зображення

Попередня обробка: Перший крок полягає в тому, щоб "очистити" відсканований документ якомога більше. Програмне забезпечення OCR використовує різні методи для видалення недоліків з низької якості сканування, вирівнювання кривих зображень і видалення будь-яких розмиття, щоб зробити процес OCR максимально точним. Крапки

видаляються, контраст регулюється, а текст загострюється. Після попередньої обробки починається OCR.[4]

Перший прохід: Більшість сучасного програмного забезпечення для розпізнавання документів працює за принципом двох проходів. Це означає, що програмне забезпечення проходить через кожен документ, який він сканує двічі. Перший "пропуск" здійснюється без будь-яких попередніх знань про документ. Програма сканує загальноприйняті символи і розбиває їх на їх основні форми, вибираючи букви з високою схожістю. [4]

Другий прохід: Наступним кроком є прийняття літер, з якими програма бачить високу схожість і використовує внутрішній словник, щоб почати порівнювати символи. Найкраще програмне забезпечення OCR для сканування документів здатне перевіряти граматику і використання у реченні, так само, як і хороший текстовий редактор. Хоча точність ніколи не є гарантією, ці два пропуски можуть забезпечити принаймні достатню точність, і методологія постійно вдосконалюється.

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		22

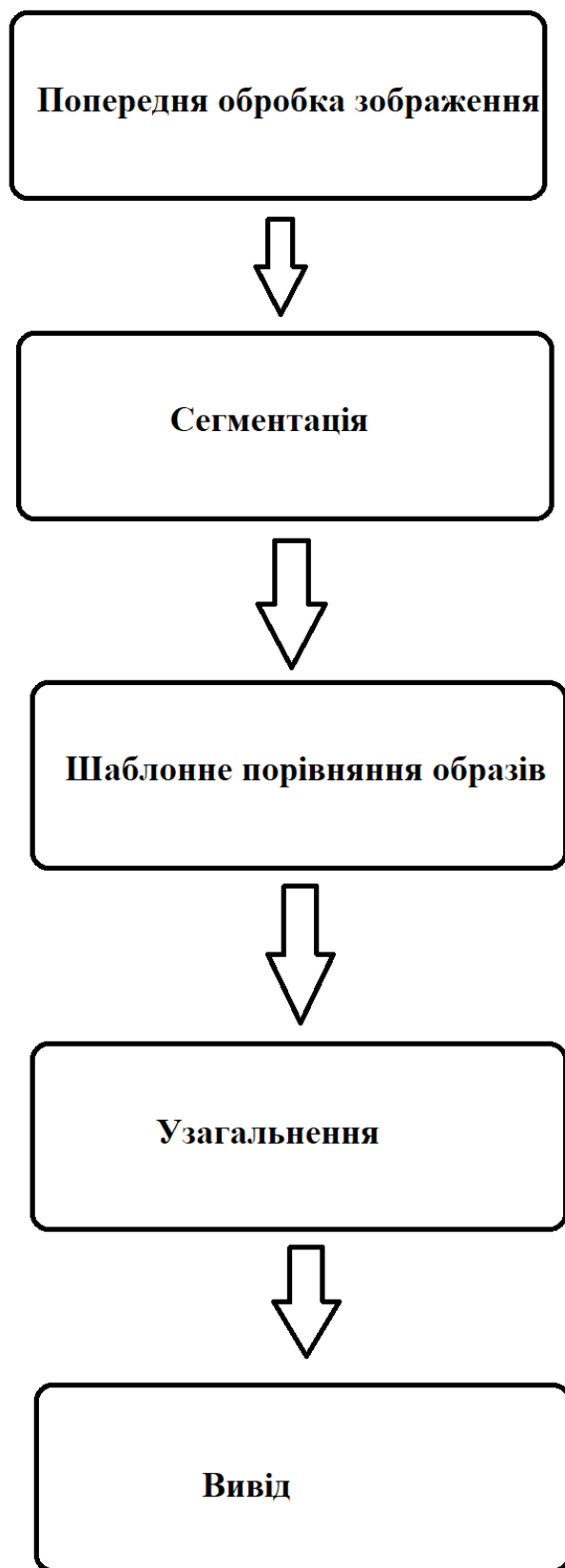


Рисунок 1 Етапи розпізнавання символу

Розпізнавання символів повинно проводитись за таким алгоритмом:

- 1) завантаження параметрів по замовчуванню під час запуску програми;
- 2) створення еталонних зображень (створивши еталони, користувач може перенавчити мережу так, щоб створений ним образ символу був еталоном);
- 3) завантаження чи створення зображення символу для розпізнавання;
- 4) розпізнавання символу системою;
- 5) виведення результату розпізнавання.

2.3 Сегментація

Одним з найпоширеніших підходів до сегментації видобутку рядків у друкованих документах є метод проектування. Він також був адаптований для рукописних документів з невеликим перекриттям. Цей простий спосіб підраховує кількість чорних пікселів у кожному рядку зображення. Ті лінії, що містять кількість пікселів нижче певного значення, вважаються лініями без 36 тексту, а інші - як текстові. Відхиленням від технології проекційного профілю є згладжений проекційний профіль, який використовує фільтр низьких частот для зменшення шуму. Інший цікавий метод, який застосовується до багатокожких рукописних документів для вилучення текстових ліній компонентів. Цей алгоритм передбачає, що гіпотетична вода тече з лівого та правого боків зображення, а символи текстових ліній працюють як дамба, яка блокує воду. Отримані ділянки, що залишилися незміцнені на зображенні, нарешті позначаються як кандидати на текстові рядки регіонів. Однак цей спосіб потребує використання структурованого елемента, який визначає кут нахилу для затоплення, який залежить від лінійних схилів кожного конкретного документа. Наприклад застосування цього методу до зразка рукописного документа. Оцінка відповідного значення параметра (кут нахилу необхідного структурованого елемента) для кожного документа є складною і визначає обчислювальні показники роботи методу. Зразок

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		24

двоетапного документа та (справа) сегментація лінії гіпотетичним потоком води. Сіримі областями є затоплені частини документа. Створення бінарної карти переходу зображення зображення для пошуку імовірного розташування текстових ліній застосовується в. Пізніше виявлені лінії очищаються за допомогою алгоритму розрізання графіків з мінімальним розрахунком / максимальним потоком. Інший підхід використовує перетворення Hough на основі блоків для виявлення та вилучення необмеженої лінії. Ніколас та інші запропонували метод сегментації ліній у документах, який базується на структурі вирішення проблем ІА виробничих систем. Лі та інші оцінюють карту вірогідності зображення документа, де кожен компонент представляє ймовірність того, що розглянутий піксель належить до текстової лінії. Вони розвивають початковий виявлення тексту та використовують метод встановлення рівня, щоб визначити межу сусідніх текстових ліній. Їхні експерименти показали, що запропонований алгоритм є надійним для масштабування зміни, обертання та шуму. У недавньому документі використовується запропонований метод вилучення текстових рядків на основі морфології для виявлення текстових областей у 37 зображеннях. Для обробки з перекосом рукописних текстових рядків автори застосовують метод на основі моментів для оцінки лінійних орієнтацій. Лікформан-Сулем та інші опублікував у 2007 році недавній оглядовий документ про текстову лінію (друковане та рукописне), сегментація історичних зображень документа. Ця робота детально описує складність цієї проблеми, а також описує класифікацію методів сегментації. Процес вилучення рядків із документа використовується як основа для вилучення структури документів, розпізнавання рукописного тексту або вдосконалення тексту. Існує безліч методів, які стосуються проблеми вилучення рядка друкованого документа, яка, як правило, скорочується до глобального пошуку косусів (текстові рядки паралельні один одному, але не обов'язково

горизонтальні). З іншого боку, при роботі з рукописним документом проблема стає більш складною лінії не паралельні один одному, однакові літери не мають однакових розмірів, текстові лінії мають літери, які поширюються на інші текстові лінії, висока текстова організація не може бути визначена (абзаци, підрозділи тощо). У будь-якому зображенні, документі чи ні, з рукописним або друкованим текстом кожен піксель може бути пов'язаний з важливістю (тобто, скільки пікселя впливає на загальне зображення).[2]

2.4. Шум

Зображення часто деградуються шумами. Шум може виникнути під час захоплення зображень, передачі тощо. Вилучення шумів є важливим завданням при обробці зображень. Загалом, результати шумозаглушення сильно впливають на якість обробки зображень. Кілька способів видалення шумів добре зарекомендували себе при обробці кольорових зображень. Характер проблеми видалення шумів залежить від типу шуму, що руйнує зображення. У сфері зменшення шумів зображення було запропоновано декілька методів 38 лінійного та нелінійного фільтрування. Лінійні фільтри не здатні ефективно усунути імпульсний шум, оскільки вони мають тенденцію до розмивання країв зображення. З іншого боку, нелінійні фільтри підходять для боротьби з імпульсним шумом. За останні кілька років виникли кілька нелінійних фільтрів на основі класичної та нечіткої техніки. Наприклад, більшість класичних фільтрів, які одночасно видаляють розмивання країв, а нечіткі фільтри мають можливість поєднувати збереження та згладжування краю. У порівнянні з іншими нелінійними методами нечіткі фільтри здатні представляти знання зрозумілим способом. У комп'ютерному природному сценічному аналізі широко використовуються кольорові зображення завдяки чудовим характеристикам над чорно-білими зображеннями. Розглянемо, наприклад,

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		26

кілька сусідніх об'єктів з різним кольором, але однакову світність. Ці об'єкти можуть бути інтерпретовані як об'єднані об'єкти на графічному рівні зображення через постійні інтенсивності, але можуть бути відбиті яскраво в кольоровому зображенні. Додана інформація про кольори висвітлює будь-яку можливу двозначність. Окрім переносної інформації про відображену енергію, кольорові зображення містять спектральну інформацію про об'єкт та надають важливі деталі для швидкого візуального пошуку, перевірки та точної класифікації об'єктів. Більше того, колір поверхні відносно інваріантний для (інтенсивності) освітленості, що відбувається через властивість кольорового сталість. Тому обробка кольорової інформації є відносно послідовною, коли умови освітлення навколишнього середовища змінюються, і забезпечує більш надійні та точні результати для сприйняття машини та аналізу природного місця.

Будь-яке зображення, отримане оптичним, електрооптичним чи електронним способом, може бути деградоване недосконалістю механізмів зондування. Потенційні деградації можуть мати місце у вигляді сенсора, шумів фотографічного зерна, розмивання (нефокусована камера, рух відносної об'єктної камери), випадкова атмосферна турбулентність тощо. Наявність шуму 39 в зображеннях являє собою незворотну втрату інформації. Кілька фільтрів були розроблені для видалення шумів зображень. Шум є результатом помилок у процесі отримання зображень, що призводить до значень пікселів, які не відповідають дійсності інтенсивності справжньої сцени. Зменшення шуму - це процес видалення шуму від сигналу. Методи зменшення шуму є концептуально дуже схожі, незалежно від оброблюваного сигналу, однак априорне знання характеристик очікуваного сигналу може означати, що реалізація цих прийомів сильно відрізняється залежно від типу сигналу. Зображення, зроблене датчиком, піддається фільтруванню різними фільтрами згладжування та отриманими зображеннями. Усі пристрої запису, як

аналогові, так і цифрові, мають риси, які роблять їх чутливими до шуму. Основна проблема обробки зображень полягає в зменшенні шуму від цифрового кольорового зображення. Удосконалення якості зображення було проблемою у всій області обробки зображень. Зображення постраждали від різних типів шумів. Шум зображення є небажаним, оскільки він погіршує якість зображення. Застосування зменшення шуму в обробці зображень є перспективним напрямом досліджень. Нечіткі технології вже застосовуються в декількох областях обробки зображень і мають численні практичні застосування. Звуковий шум, як правило, небажаний, різниця в яскравості або інформація про кольори розглядається як шум. Шум зображення може породжуватись у зерні плівок або в електронних шумах у датчику та схемі вхідного пристрою або в неминучому короткому шумі ідеального фотонного детектора. Шум зображення найбільш яскраво виражено в області зображень з низьким рівнем сигналу, такими як область тіні або під виглядом зображень. Алгоритми вилучення зображень часто приймають додатковий білий гаусовий шум (AWGN), який не залежить від фактичних значень RGB. Такі підходи не є повністю автоматичними і не можуть ефективно видалити колірний шум, вироблений сучасною цифровою камерою ПЗЗ. У цьому документі ми пропонуємо єдину схему для двох завдань: автоматичну оцінку та 40 видалення кольорового шуму з одного зображення за допомогою кусочно гладких моделей зображень. Ми вводимо функцію рівня шуму (NLF), яка є безперервною функцією, яка описує рівень шуму як функцію яскравості зображення. Потім ми оцінюємо верхню межу реальної функції шуму, встановивши нижню огиду до стандартних відхилень відмінностей зображення на одному сегменті. Для видалення кольору кольоровий шум значно зменшується шляхом проєціювання значень пікселів на лінію, яка відповідає значенням RGB у кожному сегменті. Потім побудовано гауссовське умовне довільне поле

(GCRF) для отримання нижчого чистого зображення із шумного вводу. Величезні експерименти проводяться для тестування пропонованого алгоритму, який, як показують, перевершує сучасні алгоритми стирання. Розпізнавання зображень вивчалось десятиліттями в галузі комп'ютерного зору, обробки зображень та статистичної обробки сигналів. Ця проблема не тільки забезпечує хорошу платформу для вивчення природних моделей зображень та алгоритмів розподілу сигналів, але також стає важливою частиною систем придбання цифрових зображень для підвищення якості зображень. Ці два напрямки є важливими і будуть розглянуті в цьому документі. Велика частина існуючої роботи з видалення зображення передбачає додавання білого гаусового шуму (AWGN) і видаляє шум, незалежний від каналів RGB. Однак тип і рівень шуму, який створюють цифрові камери, невідомі, якщо серія та марка камери, а також параметри камери (ISO, витримка затвора, діафрагма та вмикання / вимикання) невідомі. Наприклад, метадані формату файлів зображень (EXIF), додані до кожного зображення, можуть бути втрачені при перетворенні формату зображення та переміщенні файлу зображення. Тим часом, статистика кольорового шуму не залежить від каналів RGB через демоверсійний процес, вбудований в камери. Отже, поточні підходи, що використовуються для відміни, не є справді автоматичними і не можуть ефективно видаляти кольоровий шум. Цей факт перешкоджає практичному застосуванню методів видалення шуму до розпізнавання цифрових зображень та збільшення їх застосування. У деякому програмному забезпеченні, що заперечує зображення, користувач повинен визначити кількість рівних зображень регіонів для оцінки рівня шуму. Це спонукало нас прийняти сегментаційний підхід для автоматичного визначення рівня шуму з одного зображення. Оскільки рівень шуму залежить від яскравості зображення, ми пропонуємо оцінити верхню межу функції рівня шуму (NLF) з зображення. Зображення поділяється на кусочно-гладкі області, в яких середнє

значення має оцінка яскравості, а стандартне відхилення - переоцінка рівня шуму. Попередні функції рівня шумів вивчаються шляхом імітації процесу цифрової обробки зображень, і вони використовуються для правильної оцінки кривої, коли відсутні дані. Оскільки відокремлення сигналу та шуму від одного входу є недостатньо обмеженим, теоретично неможливо повністю відновити оригінальне зображення від шумового забрудненого спостереження. Метою вилучення зображень є якнайбільше збереження можливостей зображення при усуненні шуму. Існує ряд принципів, які ми хочемо зіставити при розробці алгоритмів розпізнавання зображень.[4]



Рисунок 2 Приклад зменшення шуму в зображенні

2.5. Математичні моделі зображення

Комп'ютерна обробка зображень можлива після перетворення сигналу зображення з безперервної форми в цифрову форму. Ефективність обробки залежить від адекватності моделі, що описує зображення,

необхідне для розробки алгоритмів обробки. У цьому випадку необхідно враховувати вплив передавальних і приймальних систем і каналу зв'язку на сигнал зображення. Модель зображення є системою функцій, що описують сутнісні характеристики зображення: функцію яскравості, що відображає зміну яскравості в площині зображення, просторові спектри і спектральну інтенсивність зображень, функції автокореляції.

Канал зображення містить оптичну систему, оптико-електричний перетворювач, пристрій аналого-цифрового перетворення (ADC) і обробку сигналу цифрового зображення. У загальному випадку безперервне зображення може бути представлено функцією з п'яти аргументів: трьох просторових координат, часу і довжини хвилі електромагнітного випромінювання.

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		31

3. ШТУЧНА НЕЙРОНА МЕРЕЖА

3.1 Штучна нейрона мережа та її види

Штучна нейронна мережа є концептуальною моделлю біологічної нейронної мережі і складається з пов'язаних між собою різним чином шарів штучних нейронів, які організовують загальну активну структуру і функціонально впливають на роботу один одного. У більшості архітектур штучних нейромереж активність нейрона визначається перетворенням зовнішнього сумарного впливу інших нейронів на даний нейрон. З моменту свого зародження технології штучних нейронних мереж розвивалися досить відокремлено від класичних методів, нерідко докорінно змінюючи уявлення про предмет і проблематику теорії машинного навчання і розпізнавання об'єктів, залишаючи значний вплив на теоретичний, термінологічний і методологічний апарати цих дисциплін.

Через деякий час після розвитку базових моделей штучних нейронних мереж, відбувся значний поділ науки про нейромережі на види топологій архітектури мереж і методи навчання мереж. У більшості архітектур штучних нейронних мереж функції активації нейронів фіксовані, а ваги синапсів є параметрами мережі. Деякі входи нейронів є зовнішніми входами сукупної мережі, а деякі виходи нейронів - виходами сукупної мережі.[7]

Приєднуючись до мережі, нейрони формують системи обробки інформації, які забезпечують ефективну адаптацію моделі до постійних змін від зовнішнього середовища. У процесі функціонування мережі вхідний сигнал перетворюється у вихідний.

Конкретний тип перетворення визначається як архітектура нейронної мережі і характеристики нейронних елементів, засоби контролю та синхронізації інформаційних потоків між нейронами. Важливим фактором

ефективності мережі є встановлення оптимальної кількості нейронів, типів взаємозв'язків між ними, а також відповідних правил передачі інформації.
[5]

При описі нейронної мережі використовуються кілька встановлених термінів, які в різних джерелах можуть мати різні інтерпретації, зокрема:

- Структура нейромережі - спосіб зв'язків нейронів у нейромережі;
- Архітектура нейромережі - структура нейромережі та типи нейронів;
- Парадигма нейромережі - спосіб навчання та використання; іноді вміщує і поняття архітектури.

На основі однієї архітектури можуть бути реалізовані різні парадигми нейромережі і навпаки.[5]

Серед відомих архітектурних рішень виділяють групу слабозв'язаних нейронних мереж (рис.3), у випадку, коли кожний нейрон мережі зв'язаний лише із сусідніми. Якщо входи кожного нейрона зв'язані з виходами усіх решта нейронів, тоді мова йде про повнозв'язані нейромережі (рис.4).

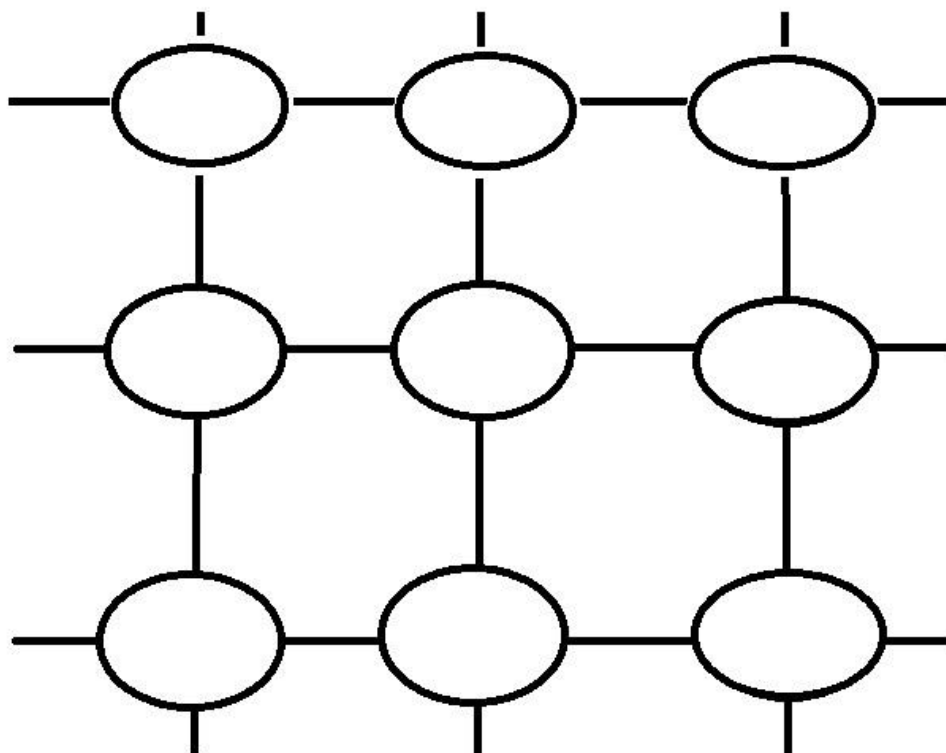


Рисунок 3 Слабозв'язна нейрона мережа

Зрозуміло, що такий поділ носить дещо теоретичний характер. Аналізуючи найбільш відомі на даний час розробки нейромереж, слід зазначити, що самим поширеним варіантом архітектури є багатошарові мережі. Нейрони в даному випадку об'єднуються у з'єднані між собою прошарки з єдиним вхідним вектором сигналів. [7]

Зовнішній вхідний вектор подається на вхідний прошарок нейронної мережі (**рецептори**). Виходами нейронної мережі є вихідні сигнали останнього прошарку (**ефектори**). Окрім вхідного та вихідного прошарків, нейромережа має один або декілька прихованих прошарків нейронів, які не мають контактів із зовнішнім середовищем. [6]

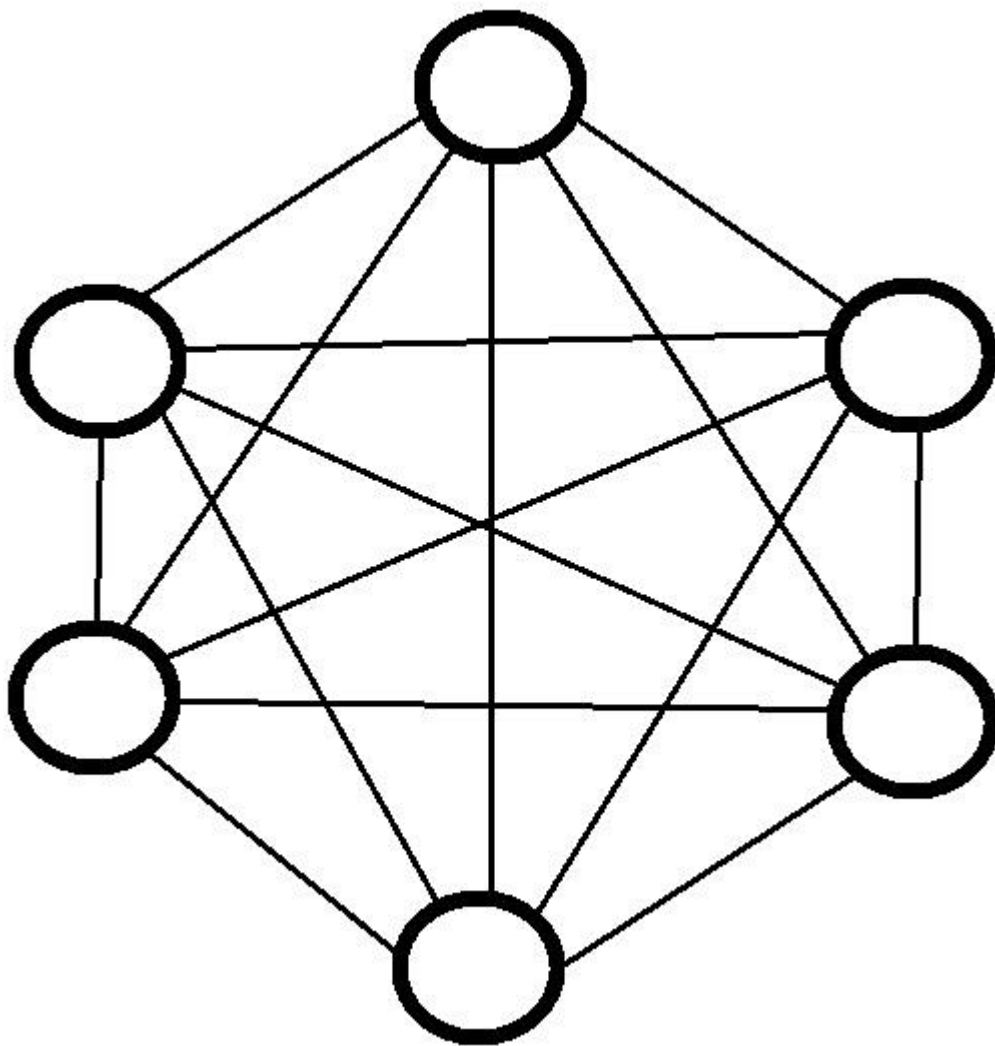


Рисунок 4 Повнозв'язні нейроні мережі

Шлях, яким нейрони з'єднуються між собою має значний вплив на роботу мережі. Більшість програмних моделей дозволяють користувачу додавати, вилучати та керувати з'єднаннями як завгодно. Корегуючи параметри, зв'язки можна робити як збуджуючими так і гальмуючими. За архітектурою зв'язків, більшість відомих нейромереж, що знайшли практичне застосування, можна згрупувати у два великих класи:

1. Мережі прямого поширення (з односкерованими послідовними зв'язками).[6]

- Перцептрони
- Мережа Back Propagation
- Мережа зустрічного поширення
- Карта Кохонена

2. Мережі зворотного поширення (з рекурентними зворотними зв'язками).

- Мережа Хопфілда
- Мережа адаптивної резонансної теорії
- Двоскерована асоціативна пам'ять

Мережі прямого поширення відносять до статичних, оскільки на задані входи нейронів надходить вектор вхідних сигналів, що незалежить від попереднього стану мережі. Рекурентні мережі вважаються динамічними, оскільки за рахунок зворотних зв'язків (петель) входи нейронів модифікуються в часі, що приводить до змін станів мережі.[5]

Оригінальність нейромереж, як аналога біологічного мозку, полягає у здібності до навчання за прикладами, що складають навчальну множину. Процес навчання нейромереж розглядається як налаштування архітектури та вагових коефіцієнтів синаптичних зв'язків відповідно до даних навчальної множини так, щоб ефективно вирішити поставлену задачу.[4]

3.2. Алгоритм роботи нейромережі

Алгоритм роботи нейронної мережі є ітеративним, його кроки називають епохами або циклами.

Епоха - одна ітерація в процесі навчання, що містить пред'явлення всіх прикладів з навчальної множини і, можливо, перевірку якості навчання на

контрольній множині.

Процес навчання здійснюється на навчальній вибірці (множині). Навчальна вибірка містить набір даних про предметну область (об'єкт, явище, процес), що поділяється на вхідні значення і відповідні їм вихідні значення набору даних.

Наприклад, після пред'явлення зображення цифри "1" на вхід нейронної мережі, вона видає деяку відповідь, не обов'язково вірну.

В навчальній множині присутня вірна (бажана) відповідь, і сенсом навчання є те, щоб на виході нейронної мережі з міткою "1" рівень сигналу був максимальний. В ході навчання нейронна мережа знаходить певні залежності вихідних полів від вхідних.

Зазвичай, в якості бажаного виходу в задачі класифікації беруть набір (1, 0, 0, ...), де 1 стоїть на виході з міткою "1", а 0 - на всіх інших виходах. Обчислюючи різницю між бажаною і реальною відповідями мережі, утворюється функція похибки. Функція похибки - це цільова функція, дозволяє оцінити якість роботи нейронної мережі і потребує мінімізації в процесі керованого навчання нейронної мережі.[4]

Алгоритмом навчання є набір формул, який дозволяє за функцією похибки обчислити необхідні поправки для вагових коефіцієнтів зв'язків нейронної мережі.

Одну і ту ж букву (а також різні зображення однієї і тієї ж букви) можна пред'являти нейронній мережі багато разів. У цьому сенсі навчання швидше нагадує повторення вправ в спорті - тренування. .[4]

Виявляється, що після багаторазового пред'явлення прикладів вагові коефіцієнти зв'язків нейронної мережі стабілізуються, причому нейронна мережа дає правильні відповіді на всі (або майже всі) приклади з бази даних. У такому випадку говорять, що "нейронна мережа вивчила всі приклади", "нейронна мережа навчена", або "нейронна мережа натренована".

Які вхідні поля (ознаки) необхідно використовувати? Спочатку вибір здійснюється евристичний, далі кількість входів може бути змінено. Складність може викликати питання про кількість прикладів в наборі даних.

Вся інформація, яку нейронна мережа має про завдання, міститься в наборі прикладів. Тому, якість навчання нейронної мережі безпосередньо залежить від кількості прикладів в навчальній вибірці, а також від того, наскільки повно ці приклади описують це завдання. Так, наприклад, безглуздо використовувати нейронну мережу для передбачення фінансової кризи, якщо в навчальній вибірці не представлено жодної кризи. Для повноцінного навчання нейронної мережі потрібна репрезентативна і доволі велика вибірка з сотні прикладів.[7]

Кількість необхідних прикладів залежить від складності розв'язуваної задачі. При збільшенні кількості ознак кількість прикладів зростає нелінійно, ця проблема носить назву "прокляття розмірності". За недостатньої кількості даних рекомендується використовувати лінійну модель.[6]

Розробник має надати можливості вибору кількості шарів у мережі і кількості нейронів в кожному шарі. Далі необхідно призначити такі значення ваг і зсувів, які зможуть мінімізувати похибку рішення.

Від якості навчання нейронної мережі залежить її здатність вирішувати поставлені перед нею завдання.

3.3. Навчання нейромережі

3.3.1 Навчання з учителем

Визначення методу навчання з учителем можна сформулювати так: задача машинного навчання у реалізації функції виведення на підставі відомих навчальних даних. Структура навчальних даних складається з набору прикладів для виконання навчання. В даному методі навчання

кожен приклад представляє собою пару, що складається з вхідного об'єкта (зазвичай вектора) і бажаного вихідного значення (контрольний сигнал). Необхідно прорахувати значення функції на момент входу даних, потім порівняти значені із значенням очікуваного результату, вирахувати помилку і скорегувати параметри (вага нейронів) мережі на цю помилку. Алгоритм навчання з учителем :

1. Взяті дані для навчання мережі розділити на дві частини. Першою половиною буде навчальна вибірка, другою - тестова вибірка. (можливо розділити дані 70/30).

2. Прорахувати значення функції для навчальної вибірки та знайти значення функції помилки.

3. Провести корегування ваг синапсів у мережі.

4. Повторювати пункти 2 і 3 до тих пір поки значення функції помилки не буде мінімальним. Повторювати ці дії можна як для кожного екземпляра навчальної вибірки так і для всієї вибірки в цілому. У першому випадку навчання буде повільніше, але з більшою точністю. У другому випадку, показник точності занепадає, але навчання буде відбуватися швидше.

5. Перевірити всі значення за тестовою вибіркою, щоб зрозуміти наскільки добре система змогла узагальнити дані (навчитися). Якщо нейронна мережа навчається з використанням заздалегідь відомих правильних відповідей, то такий алгоритм навчання називається - навчання з учителем. Необхідно відзначити, що при навчанні з учителем потрібна велика вибірка, щоб в достатній мірі сформувати робочу і гнучку нейронну мережу.

3.3.2 Навчання без учителя

Метод навчання без учителя, також називають методом неконтрольованого навчання, який має змогу знайти структуру або

відносини між різними входами. Головна відмінність від методу навчання "з вчителем", це те що в наявності є тільки вхідні дані. Алгоритм навчання без вчителя застосовується по суті тоді, коли відомі тільки вхідні дані. На основі їх мережа вчиться видавати найкращі вихідні результати. Поняттям «найкращих результатів» визначається самим алгоритмом навчання. Зазвичай алгоритм підлаштовує параметри так, щоб мережа видавала однакові результати для достатньо близьких вхідних значень.

Найбільш важливим неконтрольованим навчанням є кластеризація, яка створює різні кластери введення і зможе вносити будь-які нові дані до відповідного кластеру. Крім кластеризації є інші методи неконтрольованого навчання: виявлення аномалій, навчання і навчання в теорії Хебба. Приховані змінні моделі, такі як алгоритм максимізації очікувань, метод моментів і методи поділу сліпих сигналів. Хоча даний метод і часто використовується у прикладних задачах, він часто піддається критиці вчених через свою біологічну схильність. [4]

Важко уявити, що у мозку присутній механізм порівняння отриманих результатів з бажаними.

3.3.3 Навчання засноване на корекції помилок

Цей метод навчання перцептрона, запропонований Френком Розенблатом ще у далекому 1957 році [2], являє собою такий метод навчання, при якому вага зв'язку змінюється, поки перцептрон видає правильний результат. Параметри впроваджують зміни у випадку неспівпадання вихідних значень з вхідними. Для обчислення величини корекції використовується різниця між реальним і бажаним значенням виходу мережі. Дана модель використовує навчання з учителем, тобто навчальна множина складається із сукупності вхідних векторів для кожного з яких вказано вихідний вектор. Не зважаючи на деякі обмеження модель стала основою для багатьох сучасних більш складних алгоритмів

навчання.

3.3.4 Конкурентне навчання

Метод конкурентного навчання полягає у змаганні кожного з вихідних нейронів мережі за активацію. З цього можна зробити висновок що з усіх вихідних нейронів до роботи приступає лише нейрон із самим великим виходом. Такий алгоритм має нагадування біологічних нейронних мереж.

За допомогою методу конкуренції можна класифікувати вихідні дані де схожі приклади групуються до одного класу і подаються як один зразковий елемент. При цьому кожни нейрон із множини нейронів несе відповідаль лише за один клас. Загальне число класів з якими працює мережа дорівнюватиме число вихідних нейронів. Конкурентне навчання дуже зручно використовуваних в задачах класифікації вхідних образів. У цьому випадку кожен нейрон вихідного шару відповідає за один образ. [3]

3.3.5 Навчання Хебба

Навчання Хебба засновано на фізіологічних і психологічних дослідженнях. Алгоритм навчання можна представити у вигляді правил з 2 частин. Де перша частина звучить так: якщо два нейрона по обидві сторони синапсу активізуються одночасно, то міцність цього з'єднання зростає. І друга частина - якщо два нейрона по обидві сторони синапсу активізуються асинхронно, то такий синапс послаблюється або взагалі відмирає.[4]

3.3.6 Алгоритм зворотнього поширення помилки

Згорткові мережі за своїм типом поширення активаційного сигналу між нейронами є прямими, тому застосування алгоритму зворотного поширення помилки є цілком доречним до таких мереж. Даний алгоритм ще називають алгоритмом градієнтного спуску через те, що стратегія

підбору такого важливого параметру, як вага для кожного нейрону багатошарової мережі базується на градієтному методі. Безперервна цільова функція як показник успішності мережі в загальному випадку визначається як квадратична різниця суми між фактичним результатом і очікуваним вихідним значенням. [2]

Алгоритм зворотного поширення помилки при навчанні використовує два поширення мережею – пряме та зворотнє. На самому початку алгоритму відбувається саме прямий прохід де вхідні дані у вигляді вектору реалізують поширення між шарами, від початкових до 49 останніх.

В результаті прямого проходу генерується набір вихідних сигналів, який і визначає реакцією мережі на вхідні дані. Під час прямого проходу усі синаптичні ваги мережі є фіксованими. Другим етапом алгоритму є зворотній прохід де параметри (усі синаптичні ваги) налаштовуються відповідно за правилами корекції помилок. Суть згаданого правила така: від очікуваних вихідних значень віднімається отримане (результуюче) значення фактичного виходу і в результаті такої операції формується сигнал помилки. Сигнал помилки поширюється, як відлуння, в протилежному синаптичних зв'язків, тому алгоритм і отримав таку назву.

А синаптичні ваги у свою чергу підлаштовуються з метою максимального наближення вихідного сигналу мережі очікуваного результату. Недоліком алгоритму зворотного поширення помилки є те, що він не дозволяє в загальному випадку досягти глобального мінімуму. Тому і постають основні труднощі навчання нейронних мереж, що полягають в методах виходу з локальних мінімумів. Головними недоліками градієтного спуску або алгоритму зворотнього поширення помилки при навчанні мережі є:

- «Параліч» мережі. Значення ваг мережі в результаті корекції можуть досягти дуже великих величин. Оскільки помилка, що посилюється назад в

процесі навчання, пропорційна похідній стискаючій функції, процес навчання може майже зупинитися. Цьому можна запобігти, зменшуючи крок, але процес навчання при цьому буде відбуватися довше. [3]

- Розмір кроку. Якщо значення кроку не змінюється і воно досить мале, то метод сходиться занадто повільно. Якщо ж крок занадто великий, то може виникнути параліч мережі. Необхідно змінювати значення кроку: збільшувати до тих пір, 50 поки не припиниться поліпшення оцінки в напрямку антиградієнта і зменшувати, якщо оцінка не поліпшується. [3]

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		43

4. ТЕСТУВАННЯ СИСТЕМИ

ПЗ розпізнає прості арифметичні вирази в образах і оцінює їх, воно забезпечує зв'язок між почерком і ефективною комп'ютерною оцінкою.

Аналіз ПЗ приймає зображення письмового арифметичного виразу, потім знаходить символи на зображенні, обводить рамку для кожного символу, а потім змінює розмір зображення(рамки) кожного символу відповідно до розміру навчальних даних, після чого кожен символ класифікується як цифра (0 - 9) або операнда (+) і визначає значення символу, використовуючи штучну нейронну мережу виразу, утвореного всіма символами і відображає його користувачеві.

Модулі проектування складається з 4 компонентів:

1. I / O (main) - взаємодіє з користувачем, завантажує файли зображень, викликає функцію обробки зображень.
2. Segmenter - значення символів у зображенні.
3. Converter - змінює розміри зображень відповідно до навчальних даних (28x28).
4. Neural Network (NN) - тренує і зберігає нейронну мережу, обробляє зображення символів за допомогою NN.

Також використано OpenCV – бібліотека комп'ютерного зору, для знаходження області символу на картинці.

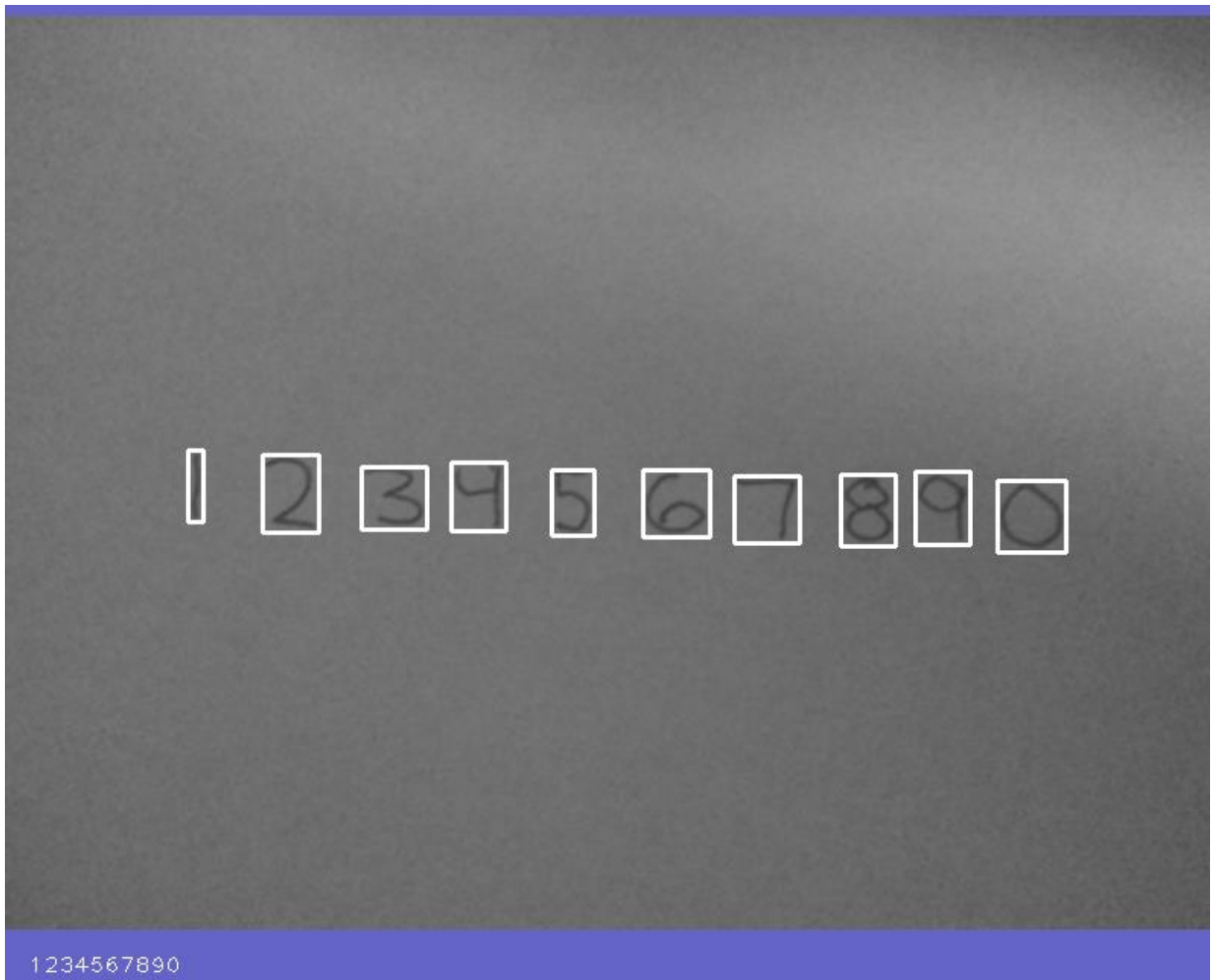


Рисунок 5 Тестування розпізнавання цифр

На рисунку 5 ми бачимо, що наша система ідеально розпізнала всі цифри. Кожна цифра була знайдена, а потім успішно порівнянна з вектором даних нейромережі.

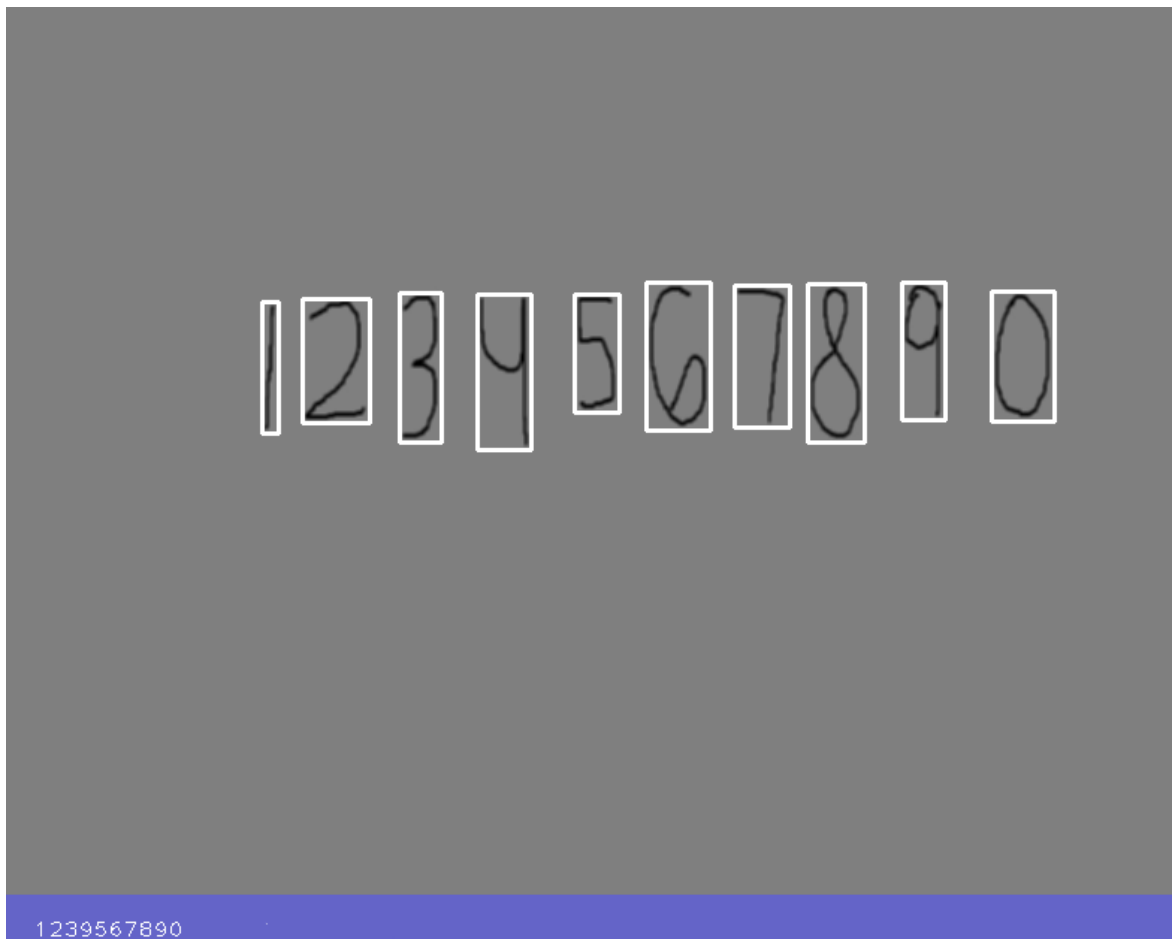


Рисунок 6 Похибка при розпізнаванні «4»

На рисунку 6 спостерігаємо, що образ «4» дуже схожий з образом «9», тому і виникла похибка. Дана похибка могла виникнути через “перенавчання”, тому проводити довге навчання нейромережі також може негативно вплинути на ефективність.

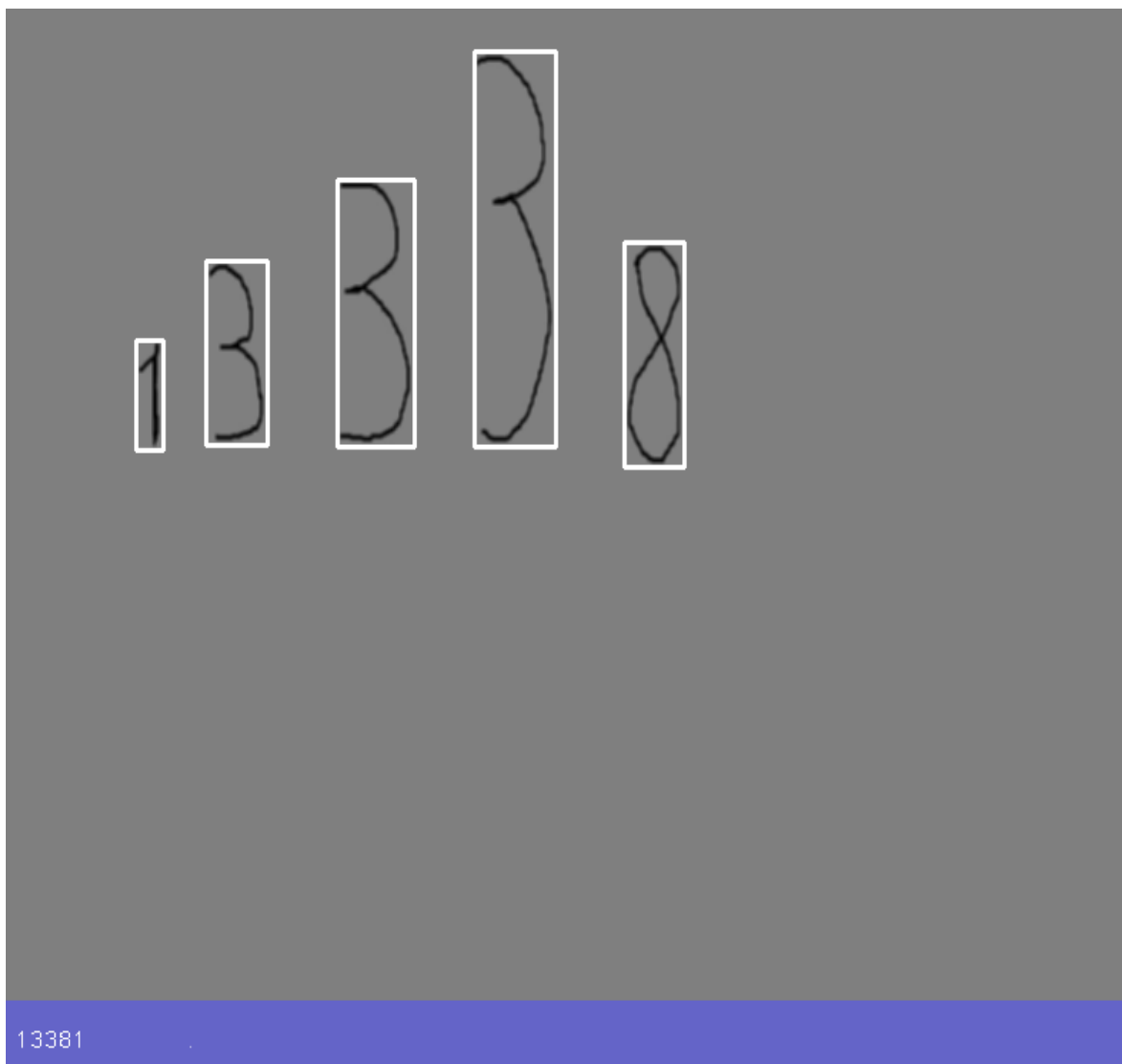


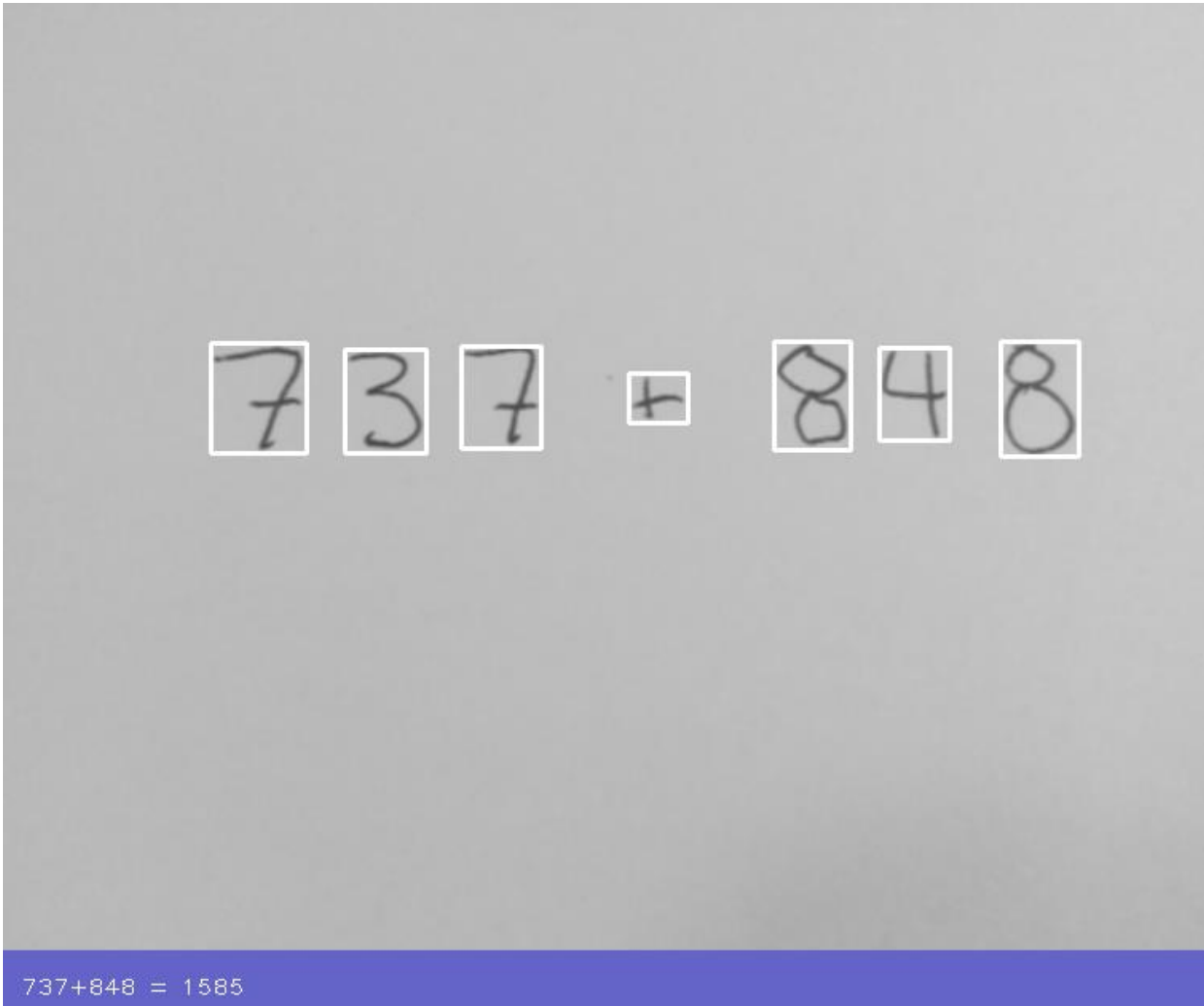
Рисунок 7 Розпізнавання символів різного розміру

На рисунку 7 система OCR успішно розпізнала образи різного розміру. Після знаходження області певного символу програмний модуль приводить всі зображення до розміру 28x28 пікселів, для порівняння їх з навчальними даними неймережі про образ символу відповідних розмірів



Рисунок 8 Арифметичні дії

На рисунку 8,9 OCR розпізнала символ «+» та виконала арифметичне додавання. При знаходженні «+» з вектора символів («5», «+», «1», «+», «2») утворюється арифметичний вираз «5+1+2» та виводиться його результат.



The image shows a handwritten arithmetic problem on a light gray background. The numbers 737 and 848 are written in a cursive style, with each digit enclosed in a small white square box. A plus sign is also in a box between the two numbers. Below the equation, a blue horizontal bar contains the result of the addition: 737+848 = 1585.

$$737 + 848 = 1585$$

Рисунок 9 Арифметичні дії

На рисунку 8 состерігаємо інший варіант (декілька цифп в числі).

ВИСНОВОКИ

В ході виконання дипломного проекту були розглянуті основні методи та алгоритми для розпізнавання образів на картинці, а також створено систему розпізнавання цифр та арифметичного додавання .

В ході розробки даної системи були розглянуті основні методи по обробці зображень та сегментації його. Використанно нейромережу для змоги розпізнавати рукописні символи.

Розроблена система має простий користувацький інтерфейс, завантаження зображення. Виділення області символу на картинці та вивід результату оптичного розпізнавання відбувається автоматично після виконання всіх модлів програми.

Навчання нейронної мережі відбувалось за допомогою обробки багатьох зображень. При спотворенних цифрах програма відносить до найбільш схожої за образом відповідної цифри. Точність при розпізнаванні 100 цифр різного почерку склала 86%, що свідчить про високу точність розпізнавання і можливість подальшого розвинення даного проекту.

					ІАЛЦ. 045490.004 ПЗ	Лист
Зм	Лист	№ докум.	Підп.	Дата		50

Література

1. Дивак М.П. Методичний посібник з дисципліни «Системний аналіз» Тернопіль: Тернопільська академія народного господарства, 2004. – 136 с.
2. B. Pang and L. Lee. Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval archive, 2008.
3. T. Mikolov et al. Efficient estimation of word representations in vector space, 2013.
4. K. S. Tai et al. Improved semantic representations from tree-structured long short-term memory network, 2015.
5. Michael A. Nielsen. Neural Networks and Deep Learning. Determination Press, 2015.
6. Ліла В.Б. Алгоритм та програмна реалізація адаптивного метода навчання штучних нейронних мереж.
7. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика, 1992.
8. Іван Гудфелов, Йоша Бенгіо, Арон Коурвілле. «Машинне навчання» MIT Press, 2016.
9. IXBT [Електронний ресурс].
10. Tesseract [Електронний ресурс].
11. ECMAScript [Електронний ресурс]. - Режим доступу до ресурсу: <https://uk.wikipedia.org/wiki/ECMAScript>